Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta 1643 (2003) 55–64
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

Nuclear localization domains in human thyroid transcription factor 2

Maria Grazia Romanelli*, Luciano Tato', Pamela Lorenzi, Carlo Morandi

Department of Mother and Child, Biology and Genetics, University of Verona, Strada le Grazie 8, 37134 Verona, Italy

Received 14 October 2002; received in revised form 8 September 2003; accepted 12 September 2003

Abstract

Thyroid transcription factor-2 (TTF2) is a nuclear protein involved in morphogenesis and gene expression in thyroid gland, belonging to the family of the *forkhead*/winged-helix transcription factors. In the present study we have investigated the sequence determinants for transport and accumulation into the nucleus of the TTF2 protein. By transient expression of fusion proteins constructed by joining different parts of TTF2 to the reporter gene of the jellyfish green fluorescent protein (GFP) and, in a separate set of deleted constructs, the glutathione *S*-transferase (GST) coding sequence, we have demonstrated that a basic amino acid stretch present at both ends of the DNA-binding domain is a bona fide nuclear localization signal (NLS). We have analyzed the subcellular localization of deleted GFP-GST-TTF2 fusion proteins and have shown that residues inside the *forkhead* domain (FHD) contributed to the complete nuclear TTF2 protein accumulation. Furthermore, by means of GST binding assays we have shown that distinct TTF2 fragments, containing the NLS, were able to bind the nuclear import receptor importin α. Taken together, our results provide the first documentation about nuclear targeting of a *forkhead* protein containing two identical NLS signal flanking the DNA-binding domain.

Keywords: Nuclear localization signal; TTF2; Forkhead; FOXE1; FKHL15; Importin α

1. Introduction

Thyroid transcription factor 2 (TTF2), also named FOXE1 (Forkhead Box E1), is a member of a gene family that encodes proteins characterized by an evolutionarily conserved 110-amino-acid DNA-binding domain, known as the *forkhead* domain (FHD) [1,2]. A unified nomenclature that uses the FOX symbol (Forkhead box) has been adopted for all chordate winged helix/*forkhead* transcription factors [3]. Many members of the family of *forkhead*/winged-helix transcription factors are key regulators of embryogenesis and play important roles in cell differentiation and development [4]. Recently, a number of mutations in human FOX genes have been associated to specific syndromes (i.e. Blepharophimosis/Ptosis/Epicanthus Inversus Syndrome [5]; Axenfeld–Rieger syndrome [6]; speech disorders [7], anterior segment ocular dysgenesis and cataracts [8]. TTF2 has been associated to thyroid agenesis, cleft palate and choanal atresia [9].

Efficient nuclear localization of transcription factors is essential for their function [10]. Nuclear proteins may accumulate in the cell nucleus selectively and efficiently by active transport mediated by specific nuclear localization signal (NLS) recognized by cytosolic receptors [11,12]. The canonical NLSs are characterized by short amino acid stretches that are enriched in basic amino acid and may be sufficient to confer nuclear localization when they are conjugated to a carrier protein. The majority of NLSs hitherto known fall into two well-defined classes, the first one characterized by a short stretch of basic amino acids, exemplified by the SV40 large T antigen NLS (T-ag P<u>*KKKRK*</u>V), and the second one characterized by a bipartite NLS, composed of two stretches of basic amino acids separated by a spacer of 10–12 amino acids, exemplified by nucleoplasmin (<u>*KR*</u>PAATKKAGQA<u>*KKKK*</u>) [13]. The proteins containing the "classical" NLSs are imported into the nucleus by the importins α/βheterodimers. The α importins contain the NLS binding site, whereas importins-β mediate the translocation through the nuclear pore [14,15]. Highly homologous signal sequences may be imported via a different import pathway [16].

In order to identify the functional domains required for TTF2 nuclear localization, we have investigated the subcel-

* Corresponding author. Tel.: +39-45-802-7182; fax: +39-45-802-7180.
E-mail address: mariagrazia.romanelli@univr.it (M.G. Romanelli).

0167-4889/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamcr.2003.09.002

lular distribution of different GFP-tagged deletion mutants of TTF2. We have shown that a basic residues stretch present at the boundaries of the FHD is a functional NLS required for cytosolic receptor recognition.

2. Materials and methods

2.1. Bioinformatic analyses

Amino acid sequences of FOX proteins were retrieved using BLAST search program at the server supported through NCBI (http://www.ncbi.nlm.nih.gov/blast/). FHD sequences from human FOX proteins shown in Fig. 1 were aligned using ClustalW, accessed through the EMBL European Bioinformatic Institute (http://www2.ebi.ac.uk/clustalw/).

2.2. Construction of expression vectors

DNA manipulations and cloning experiments were performed according to standard procedures [17]. The GFP constructs were obtained by inserting either the sequence coding for the entire human TTF2 protein or parts of it into the *Sal*I–*Bam*HI cleaved pEGFP-C1 (enhanced green fluorescent protein C-terminal protein fusion vector; Clontech Laboratories). PGFP-GST construct contains the full-length glutathione *S*-transferase (GST) coding sequence cloned in frame into the *Eco*RI–*Sal*I cleaved pEGFP-C1 vector. The GST sequence was polymerase chain reaction (PCR)-amplified from pGEX-5X-1 vector with forward primer 5′-CCGGAATTCTCCCC-TATACTAGGTTATTG-3′and reverse primer 5′-CAGG-CATTCAGTCGACTCAGTCACGATGCGGCCGCT-3′. The RRRKR motif or portion of TTF2 DNA-binding domain (residues 1 to 80, 1 to 110, 80 to 110, 80 to 151, and 110 to 151) was cloned separately into the *Sal*I–*Bam*HI cleaved pGFP-GST expression vector. The TTF2 sequences were obtained from the FKHL15 cDNA [18] by PCR using the primers described below. Each pair contains a *Sal*I restriction site (forward primers, F) or a *Bam*HI restriction site (reverse primers, R) upstream to the sequence matching the human TTF2 coding sequence (EMBL/GenBank accession number: U89995). F1: 5′-GGCAGGCATTCA*GTCGAC* ATGACTGCCGAGAGCGGG-3′; F2; 5′-GGAAGA-CATTCA*GTCGAC*CGCGGGAAGCCGCCCTAC-3′; F3: 5′-GGAAGACATTCA*GTCGAC* AGCTTCCTGCG-CCGCCGC-3′; F4: 5′-GGAAGACATTCA*GTCGAC*AAGTT-CATCACCGAGCGCTTCCCCTTCTAC-3′; F5: 5′-GGAAGACATTCA*GTCGAC*AAGTTCATCACC-GAGCGCTTCCCCTTCTAC-3′; F6: 5′-GGAAGACA-TTCA*GTCGAC* CGGCGCCGCAAGCGC-3′; R1: 5′-CGC*GGATCC* TCACATGGCGGACACGAACCG-3′; R2: 5′-ATT*GGATCC* TCAGCGCTTGAAGCGCTT-3′; R3: 5′-ATT*GGATCC* TCACTGCAGGGGGCGCTT-3′; R4 GGCAGGC*ATTCAGTCGAC* TCAGTCACGATGCGGCC-GCT; R5: 5′GGCAGGC*ATTCAGTCGAC* TCAGCGCTT-GCGGCGCCG-3′. Underlined nucleotides represent the restriction sites. Deletions of the amino acid residues at position 43–45 or 144–146 were obtained in two steps PCR using primer R6: 5′-ATCCTCACTGCAGGGGGCGC-TTCCCGCCCGCCCCGCGGCCCG TGGCCTCCCC-TGGGA-3′, or primer R7 5′-CCTCAGCGCTTGAA-GCGCTTCAGGAAGCTG CCGCTCTCGA ACATGT CCTCCGCGTT-3′, respectively. The TTF2 deleted fragments were amplified in the first PCR and the restriction sites suitable for cloning were introduced in the second PCR using primers F1 and R2, or F1 and R3. Deletions in the FHD were obtained by recombinant PCR approach using the following primers: F 5′-ATCGCGCT-CATCGCCATCGCGCACGCG-3′and R 5′-CGCG-TGCGCGATGGCGATGAGCGCGAT-3′ for Δ 64–65; F 5′-CACGCGCCCGAGCTCACGCTGGGC-3′and R 5′-GCCCAGGTCGAGCTCGGGCGCGTG-3′ for Δ 72–73; F 5′-GAGCGCCGCCTCGGCGGCATCTACAAG-3′ and 5′-CTTGTAGATGCCGCCGAGGCGGCGCTC-3′for Δ 75–76; F 5′-TACCGCGACAACCCCTGGCAGAACAGCATC-3′ and R 5′-GATGCTGTTCTGCCAG-GGGTTGTCGCGGTA-3′ for Δ 95–96; F 5′-TG-GCAGAACAGCATCAACCTCACACTCAAC-3′ and R5′-GTTGAGTGTGAGGTTGATGCTGTTCTGCCA-3′ for Δ 102–103; F 5′-TTCCTCAAGATCCCGGC-CGGCCGCCCG-3′ and R 5′-CGGGCGGCCGGC-CGGGATCTTGAGGAA-3′ for Δ 116–117; F 5′-CTCGACCCCAACGCGATGTTCGAGAGCGGC-3′ and RGCCGCTCTCGAACATCGCGTTGGGGTCGAG-3′ for Δ 134–135 (numbers indicate the amino acid residues deleted). PGEX-TTF2 (1–151), pGEX-TTF2 (1–50), pGEX-TTF2 (51–151) and pGEX-TTF2 (51–151) Δ 144–146 constructs were obtained by inserting the TTF2 sequences, obtained from FKHL15 cDNA by PCR amplification with primers that introduced the *Eco*RI and *Hin*dIII sites at the fragments ends, into *Eco*RI–*Hin*dIII cleaved pGEX-5X-1 vector (Amersham Pharmacia Biotech) modified by introduction of a *Hin*dIII site.

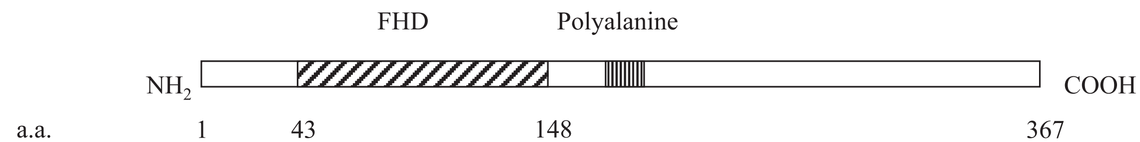
2.3. Cell culture and transfection

Adherent HeLa cells were maintained in exponential growth in Dulbecco's modified Eagles medium supplemented with 10% fetal calf serum. Cells were grown on glass coverslips and transfected using Lipofectamine (Life Technologies) according to manufacturer's instructions. Forty-eight hours after transfection, cells were washed with PBS and fixed in 2% paraformaldehyde in PBS at room temperature for 20 min. Cellular localization of the GFP-TTF2 fusion proteins was examined under a Leitz Orthoplan microscope with an epifluorescence attachment.

2.4. Analysis of chimera proteins by Western blotting

Whole cells extracts were harvested in sample buffer (50 mM Tris–HCl pH 6.8, 2% SDS, 100 mM dithiothreitol,

A



B

```
FOXE1/TTF2             RRRKRPLQRGKPPYSYIALIAMAIAHAPERRLTLGGIYKFITERFPFYRDNPK-----KWQNSIRHNLTLNDCFLKIPREAG-RPGKGNYWALPNAEDMFESGSF-LRRRKRF
FOXA1/HNF-3A/P55317    KTFKRSYPHAKPPYSYISLITMAIQRAPSKMLTLSEIYQWIMDLFPYYRQNQQ-----RWQNSIRHSLSFNDCFVKVARSPD-KPGKGSYWTLHPDSGNMFENGCYLRRQKRF
FOXA2/HNF-3B/Q9Y261    KTYRRSYTHAKPPYSYISLITMAIQQSPNKMLTLSEIYQWIMDLFPFYRQNQQ-----RWQNSIRHSLSFNDCFLKVPRSPD-KPGKGSFWTLHPDSGNMFENGCYLRRQKRF
FOXA3/HNF-3G/P55318    KGYRRPLAHAKPPYSYISLITMAIQQAPGKMLTLSEIYQWIMDLFPYYRENQQ-----RWQNSIRHSLSFNDCFVKVARSPD-KPGKGSYWALHPSSGNMFENGCYLRRQKRF
FOXB1/FKH5/AAC62493    RPGRNTYSDQKPPYSYISLTAMAIQSSPEKMLPLSEIYKFIMDRFPYYRENTQ-----RWQNSLRHNLSFNDCFIKIPRRPD-QPGKGSFWALHPSCGDMFENGSFLRRCKRF
FOXC1/FKHL7/AAC18081   PQPKDMV---KPPYSYIALITMAIQNAPDKKITLNGIYQFIMDRFPFYRDNKQ-----GWQNSIRHNLSLNECFVKVPRDDK-KPGKGSYWTLDPDSYNMFENGSFLRRRRRF
FOXC2/MFH-1/Q99958     AAPKDLV   KPPYSYIALITMAIQNAPEKKITLNGIYQFIMDRFPFYRENKQ-----GWQNSIRHNLSLNECFVKVPRDDK-KPGKGNYWTLDPQSEDMFDNGSFLRRRKRF
FOXD1/FREAC-4/Q16676   SGAKNPLV--KPPYSYIALITMAILQSPKKRLTLSEICEFISGRFPYYREKFP-----AWQNSIRHNLSLNDCFVKIPREPG-NPGKGNYWTLDPESADMFDNGSFLRRRKRF
FOXD2/FREAC-9/AAC15421 AATRSPLV--KPPYSYIALITMAILQSPKKRLTLSEICEFISGRFPYYREKFP-----AWQNSIRHNLSLNDCFVKIPREPG-NPGKGNYWTLDPESADMFDNGSFLRRRKRF
FOXD3/fkh/NP_036315    SKPKNSLV--KPPYSYIALITMAILQSPQKKLTLSGICEFISNRFPYYREKFP-----AWQNSIRHNLSLNDCFVKIPREPG-NPGKGNYWSLDPASQDMFDNGSFLRRRKRF
FOXD4/FREAC-5/Q12950   --ARQPA---KPPSSYIALITMAILQSPHKRLTLSGICAFISDRFPYYRRKFP-----AWQNSIRHNLSLNDCFVKIPREPG-RPGKGNYWSLDPASQDMFDNGSFLRRRKRF
FOXE2/HFKH4/CAA64246   GAPAPPAAR-KPPYSYIRRHAMAIG-SP--RLTLGGIYKFITEGFPFYPDNPK-----KWQNSIRHNLTINDCFLKIPREAG-RRRKGNYWALDPNAEDMFESGSFLRRRKAS
F0XE3/FREAC8/Q13461    RRRRRPLQRGKPPYSYIALIAMALAHAPGRRLTLAAIYRFITERFAFYRDSPR-----KWQNSIRHNLTLNDCFVKVPREPG-NPGKGNYWTLDPAAADMFDNGSFLRRRKRF
FOXF1/FREAC-1/Q12946   KTNAGIRRPEKPPYSYIALIVMAIQSSPTKRLTLSEIYQFLQSRFPFFRGSYQ-----GWKNSVRHNLSLNECFIKLPKGLG-RPGKGHYWTIDPASEFMFEEGSFRRRPRGF
FOXF2/FREAC-2/AAC32226 KASSGLRRPEKPPYSYIALIVMAIQSSPSKRLTLSEIYQFLQARFPFFRGAYQ-----GWKNSVRHNLSLNECFIKLPKGLG-RPGKGHYWTIDPASEFMFEEGSFRRRPRGF
FOXGI/BF2/P55316       EKKNGKYEKPPFSY--NALIMMAMRQSPEKRLTLNGIYEFIMKNFPYYRENKQ-----GWQNSIRHNLSLNKCFVKVPRHYD-DPGKGNYWMLDPSSDDVFIGGTTGKLRRST
FOXG2/BF1/P55315       EGEKKNGKYEKPPFSYNALIMMAIRQSPEKRLTLNGIYEFIMKNFPYYRENKQ-----GWQNSIRHNLSLNKCFVKVPRHYD-DPGKGNYWMLDPSSDDVFIGGTTGKLRRRS
FOXH1/FAST1/XP_005149  RRKKRYLRHDKPPYTYLAMIALVIQAAPSRRLKLAQIIRQVQAVFPFFREDYE-----GWKDSIRHNLSSNRCFRKVPKDPAKPQAKGNFWAVDVSLIPAEALRLQNTALCRR
FOXI1/HFHE/AAB50574    PSQEELMKLVRPPYSYSALIAMAIHGAPDKRLTLSQIYQYVADNFPFYNKSKA-----GWQNSIRHNLSLNDCFKKVPRDED-DPGKGNYWTLDPNCEKMFDNGNFRRKRKRK
FOXJ1/HFH4/CAA67730    DVDYATNPHVKPPYSYATLICMAMQASKATKITLSAIYKWITDNFCYFRHADP-----TWQNSIRHNLSLNKCFIKVPREKD-EPGKGGFWRIDPQYAERLLSGAFKKRRLPP
FOXK1/ILF1/NP_004505   SGGDSPKDDSKPPYSYAQLIVQAITMAPDKQLTLNGIYTHITKNYPYYRTADK-----GWQNSIRHNLSLNRYFIKVPRSQE-EPGKGSFWRIDPASESKLIEQAFRKRRPRG
FOXL1/FREAC-7/Q12952   GRAETPQK---PPYSYIALIAMAIQDAPEQRVTLNGIYQFIMDRFPFYHDNRQ-----GWQNSIRHNLSLNDCFVKVPREKG-RPGKGNYWTLDPASEDMFDNGSFLRRRKRF
FOXL2/                 TAPEKPDPAQKPPYSYVALIAMAIRESAEKRLTLSGIYQYIIAKFPFYEKNKK-----GWQNSIRHNLSLNECFIKVPREGG-GERKGNYWTLDPACEDMFEKGNYRRRRRMK
FOXM1A/HFH11A/AAC51128 SASWQNSVSERPPYSYMAMIQFAINSTERKRMTLKDIYTWIEDHFPYFKHIAKP----GWKNSIRHNLSLHDMFV---RETS-ANGKVSFWTIHPSANRYLTLDQVFKPLDPG
FOXN2/HTLF/P32314      RTQKKKSATSKPPYSFSLLIYMAIEHSPNKCLPVKEIYSWILDHFPYFATAPT-----GWKNSVRHNLSLNKCFQKVERSHGKVNGKGSLWCVDPEYKPNLIQALKKQPFSSA
FOXN3/CHES1/O00409     PYDARQNPNCKPPYSFSCLIFMAIEDSPTKRLPVKDIYNWILEHFPYFANAPT-----GWKNSVRHNLSLNKCFKKVDKERSQSIGKGSLWCIDPEYRQNLIQALKKTPYHPH
FOXO1/ALV/NP_06271     -----SRRNAWGNLSYADLITKAIESSAEKRLTLSQIYEWMVKSVPYFKDKGDSNSSAGWKNSIRHNLSLHSKFIRVQNEGT---GKSSWWMLNPEGGKSGKSPRRRAASMDN
FOXO4/AFX1/P98177      -----SRRNAWGNQSYAELISQAIESAPEKRLTLAQIYEWMVRTVPYFKDKGDSNSSAGWKNSIRHNLSLHSKFIKVHNEAT---GKSSWWMLNPEGGKSGKAPRRRAASMDN

                                     ::      .:  :    : :  *   :     ::           *::*:**.*: :  *    .      :* : *
```

Helix 1 Helix 2 Helix 3 Wing 1 Wing 2

Fig. 1. Panel A: Diagrammatic representation of the functional domains of human TTF2 protein. The TTF2 structural domains are diagrammed with boxes representing *forkhead* DNA-binding domain (FHD) and the polyalanine stretch. Panel B: Alignment of the TTF2 *forkhead* DNA-binding domain with selected human FOX representatives. Alignment was performed by ClustalW Service at the European Bioinformatic Institute. At the bottom, the consensus line indicates '*' for identical or conserved residues in all sequences; ':' for conserved substitutions; and '.' for semi-conserved substitutions. The position of predicted 'helix' and 'wing' segments are indicated at the bottom of the panel.

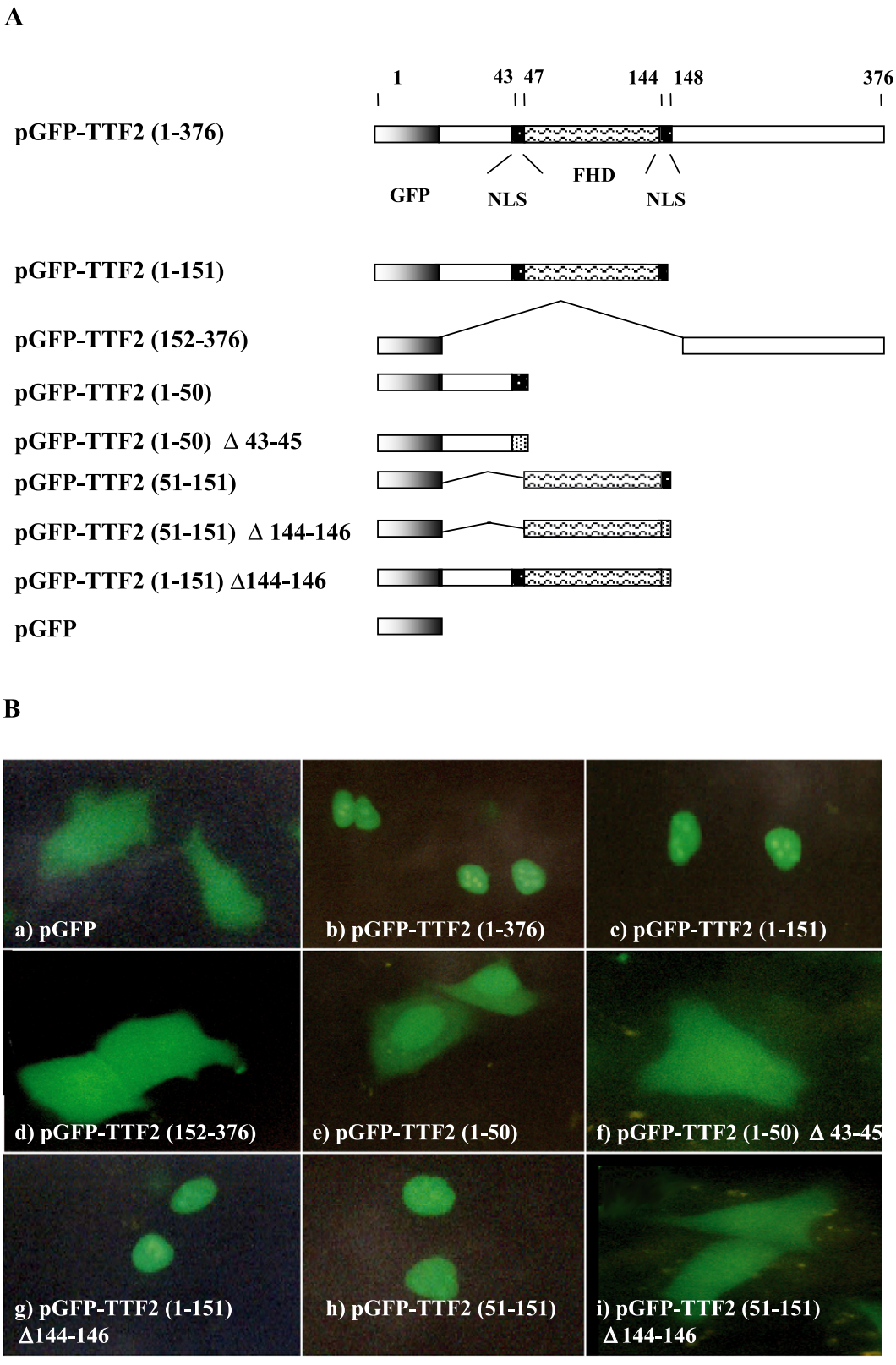


Fig. 2. Panel A: Schematic representation of expression vectors used to produces GFP-TTF2 fusion proteins. The hatched box corresponds to the FHD; the ■ black boxes correspond to the basic stretch RRRKR; deletions of basic residues are indicated with dotted boxes ⊡; shaded boxes correspond to the GFP. Numbers refer to TTF2 amino acids fused to GFP, or deleted (Δ). Panel B: The basic stretches at the boundaries of the TTF2 FHD are NLS. Fluorescence micrographs of HeLa cells transfected with plasmids expressing GFP or fusions between GFP and parts of TTF2. Numbers refer to TTF2 amino acids fused to GFP, or deleted (Δ).

10% glycerol, 0,1% bromophenol blue) from the transiently transfected HeLa cells 48 h after transfection. The extracts were subjected to 12% SDS-polyacrylamide gel electrophoresis and analyzed by Western blotting using the rabbit polyclonal GFP Antiserum (Invitrogen Life Technologies). After incubation with the anti rabbit IgG alkaline phosphatase conjugated secondary antibody (Boehringer Mannheim), the proteins were visualized using a chromogenic detection system (Boehringer Mannheim).

2.5. Expression of fusion proteins in Escherichia coli

The GST fusion system was used to generate chimeric proteins. The pGEX-TTF2 (1–151), pGEX-TTF2 (1–50), pGEX-TTF2 (51–151) and pGEX-TTF2 (51–151) Δ 144–146 constructs were introduced into the *E. coli* strain BL21 (DE3). Cell culture and batch purification of the GST fusion proteins were performed essentially according to the manufacturer's instruction (Amersham Pharmacia Biotech). All the procedures were carried out at 4 °C; 1 mM EGTA and 2 mM dithiothreitol (DTT) were included in the buffers throughout the purification procedures. Recombinant GST proteins, after elution, were dialyzed at 4 °C against binding buffer (20 mM Hepes, 150 mM KOAc, 2 mM Mg $(OAc)_2$, 2mM DTT). The concentration of proteins was determined by the Bradford's method using the Bio-Rad dye reagent (Bio-Rad) and BSA as standard. Protein samples were aliquoted, quick frozen in liquid N_2, and stored at −80 °C.

The *E. coli* strain BLR containing GST fusions of a functional SV40 large T antigen nuclear localization signal (Tag NLS) or an inverse version of Tag NLS (Tag NLSinv) was cultured, and the proteins were purified as described [19].

2.6. In vitro binding assay

The plasmid pRSET-hSRP1 containing human importin α cDNA [20] was used to produce a [^{35}S]-methionine-labeled protein by means of the Promega TNT T7 Quick Coupled Translation System. In vitro binding assays were performed essentially as described [21]. Briefly, 15 μg of GST-TTF2 (1–151), GST-TTF2 (1–50), GST-TTF2 (51–151) and GST-TTF2 (51–151) Δ 144–146 and 7 μg of GST-Tag NLS or GST-TagNLSinv were incubated with 40 μl of glutathione-agarose beads (Amersham Pharmacia Biotech) in 0.5 ml of binding buffer (20 mM HEPES, pH 6.8, 150 mM KOAc, 2 mM Mg $(OAc)_2$, 2 mM DTT, 0.1% Tween 20) for 2 h at 4 °C. One-twentieth of the beads was removed and the amount of immobilized GST fusion proteins was analyzed by SDS-PAGE. The rest of the beads was incubated with 90 μl of in vitro translated importin α reaction mixture for 4 h at 4 °C, then washed in binding buffer and boiled in 30 μl of sample buffer; the immobilized proteins were resolved on a SDS-10% polyacrylamide gel. The ^{35}S-labeled importin α bound to the GST fusion proteins was detected by fluorography using Amplify Reagent (Amersham Pharmacia Biotech).

3. Results and discussion

3.1. The basic motif at the N- and C-terminal ends of the TTF2 FHD is a functional NLS

The TTF2 gene sequence, located on chromosome 9q22, is organized in one exon and encodes for a protein of 376 amino acids. TTF2 cDNAs have been isolated from a cDNA library enriched for transcripts from chromosome 9q22 (identified as FKHL15, forkhead-like gene 15) [18], and from a human genomic library [22].

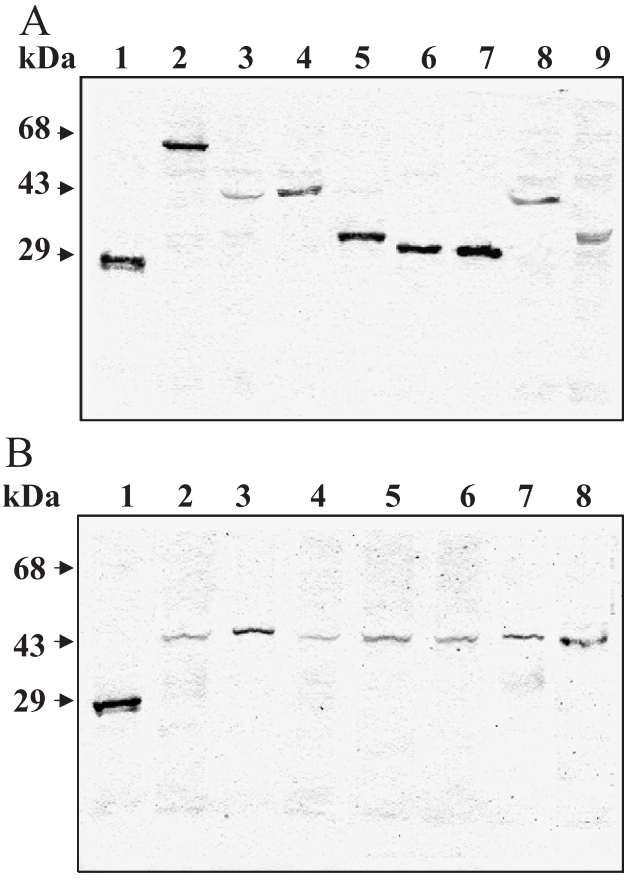


Fig. 3. Western blot analysis of GFP-TTF-2 fusion proteins transfected in HeLa cells. HeLa whole cell extracts (10 μg/lane) were separated by 12% SDS-polyacrylamide gel electrophoresis and transferred to nitrocellulose membrane. The membrane was probed with polyclonal anti-GFP antiserum. Panel A: Lane 1, GFP alone; lane 2, GFP-TTF2 (1–376); lane3, GFP-TTF2 (1–151); lane 4, GFP-TTF2 (152–376); lane 5, GFP-TTF2 (51–151) Δ 144–146; lane 6, GFP-TTF2 (1–50); lane 7, GFP-TTF2 (1–50) Δ 43–45; lane 8, GFP-TTF2 (1–151) Δ 144–146; lane 9, GFP-TTF2 (51–151). Molecular weight markers are shown on the left. Panel B: Lane 1, GFP alone; lane 2, GFP-TTF2 (1–151) ΔMA; lane 3, GFP-TTF2 (1–151) ΔKK; lane 4, GFP-TTF2 (1–151) ΔRH; lane 5, GFP-TTF2 (1–151) ΔRR; lane 6, GFP-TTF2 (1–151) ΔTL; lane 7, GFP-TTF2 (1–151) ΔRE; lane 8, GFP-TTF2 (1–151) ΔED. Molecular weight markers are shown on the left.

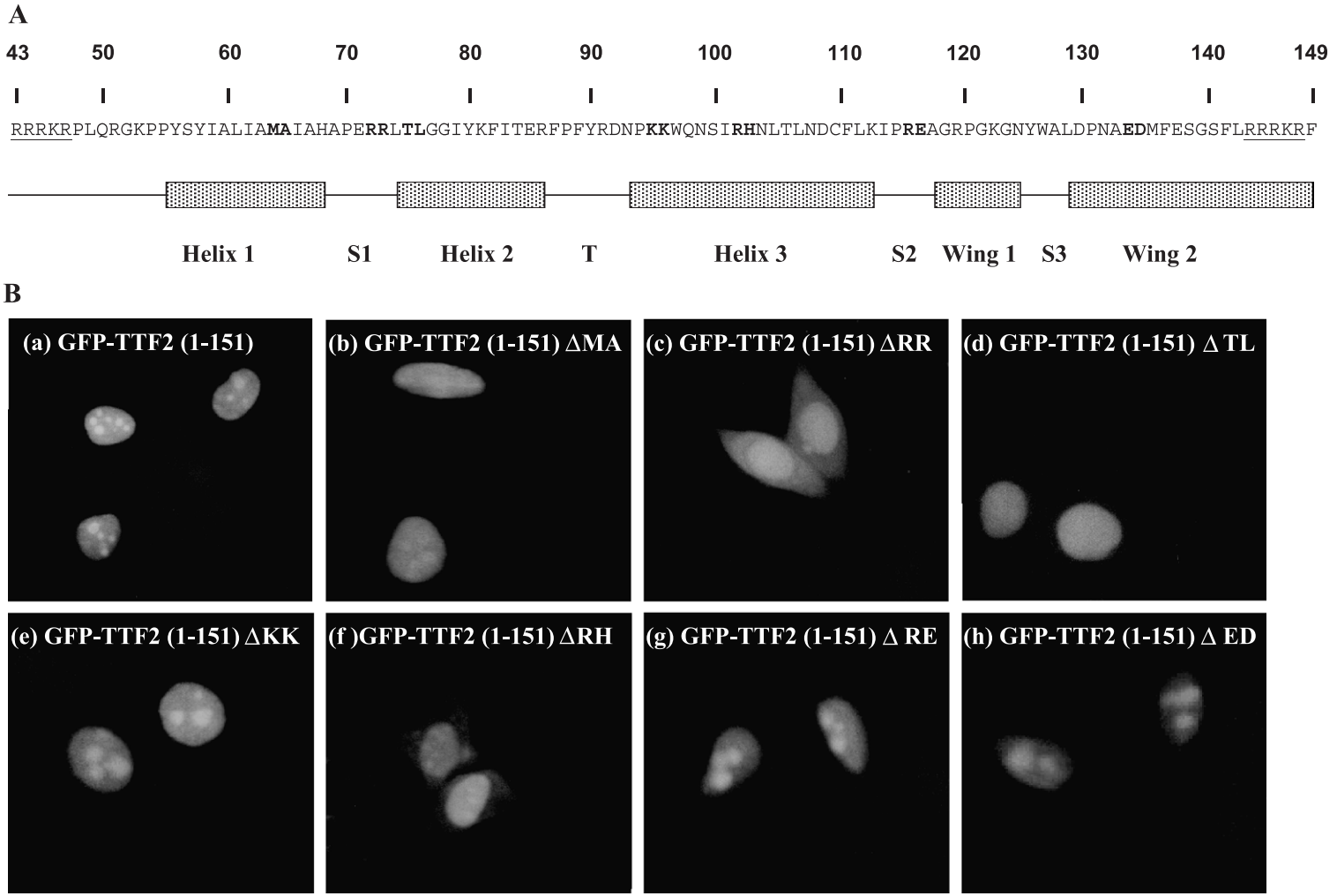


Fig. 4. Panel A: The amino acid sequence of the TTF2 FHD. Residues underlined indicate NLS stretches. Residues in bold were deleted in the fusion proteins analyzed by transfections, as shown in panel B. The positions of predicted 'helix' and 'wing' segments are indicated below the amino acid sequence. S1, S2 and S3 indicate extension of the β strands. T denotes turn region. Panel B: Residues in the FHD contribute to GFP-TTF2 (1–151) full nuclear accumulation. Fluorescence micrographs of HeLa cells transfected with plasmids expressing GFP-TTF2 (51–151) or GFP-TTF2 (1–151) fusion proteins. Letters close to Δ indicate deleted residues.

Analysis of the TTF2 amino acid sequence identifies, in addition to the FHD, a second domain rich in alanine residues that has been found in other developmental DNA binding proteins responsible for transcriptional repression activity [23–26] and two short stretches of basic amino acids (RRRKR) at both ends of the FHD (Fig. 1, panel A). Sequence alignments of representative human FOX proteins segments, containing the FHD and its N- and C-terminal boundaries, show that a stretch of basic amino acids is present in most of the proteins at the C-terminal of DNA-binding domain (Fig. 1, panel B), whereas a basic stretch at the N-terminal of the DNA-binding domain is less conserved (i.e. FOXE3 and FOXH1, other than TTF2). Previous studies have demonstrated that the basic stretches at the C-terminal of the FHD are bona fide NLSs in four FOX proteins [27–30]. These sequences differ from the basic stretch present at both ends of TTF2 FHD.

In order to analyze the contribution to nuclear localization of the N- and C-terminal ends of TTF2 FHD, we produced recombinant vectors (Fig. 2A) expressing the complete TTF2 protein, or parts thereof, individually fused to the C-terminal of the GFP, which usually diffuses to the nucleus (Fig. 2B, a). The pGFP-TTF2 vector expresses a chimeric protein of approximately 70 kDa that may not pass through the nuclear pore complex (NPC) passively. In transfected HeLa cells the full-length GFP-TTF2 (1–376) localized exclusively into the nucleus (Fig. 2B, b), indicating an active transport of the chimeric protein. We analyzed the contribution to nuclear accumulation of the TTF2 segment spanning residues 1–151, which contain the FHD and its basic boundaries, compared to that of the fragment 152 up to 376. Expression of GFP-TTF2 (1–151) showed a strictly nuclear localization (Fig. 2B, c), whereas a fusion protein in which the C-terminal residues 152 to 376 were present produced a diffuse fluorescence distributed both in the cytoplasm and nucleus (Fig. 2B, d). We then analyzed the contribution of each putative NLS to the nuclear transport of GFP-TTF2 fused proteins. GFP-TTF2 (1–50), accumulated predominantly within the nucleus of transfected cells, even if not exclusively (Fig. 2B, e), whereas deletion of residues 43–45 in the putative NLS produced a signal distributed equally in the cytoplasm and the nucleus (Fig. 2B, f). When we analyzed the cellular distribution of a chimeric GFP-TTF2 (1–151) protein, which was deleted of residues 144–146 in the second putative NLS, the protein displayed a nuclear localization (Fig. 2B, g), similar to that observed with the expression of the chimeric protein GFP-TTF2 (51–151) (Fig. 2B, h). Deletion of residues 144–146 in the GFP-TTF2 (51–151) abolished the strictly nuclear localization (Fig. 2B, i). Taken together, these results indicate that the basic stretches at the N- and C-terminal ends of the FHD contribute to nuclear translocation. The only construct containing the basic stretch that did not accumulated completely into the nucleus was the smaller fusion protein GFP-TTF2 (1–50). Western blotting using anti-GFP antibody have shown an appropriate full-length expression of all of the GFP-TTF-2 fusion

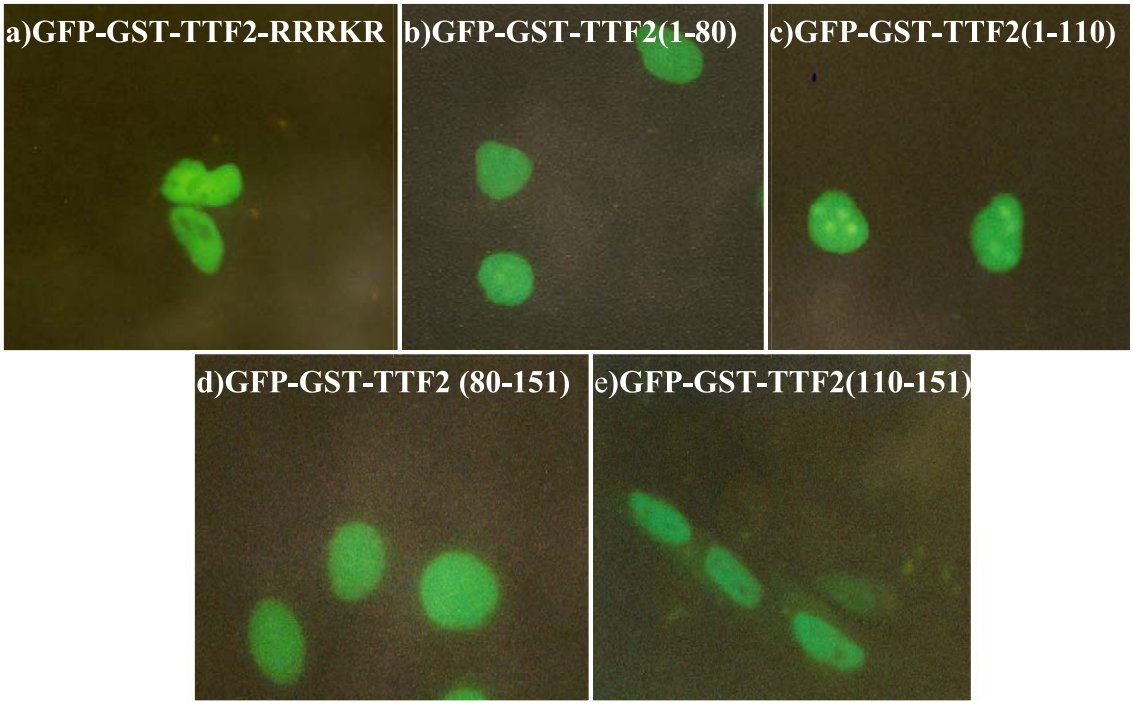


Fig. 5. RRRKR is a bona fide NLS. Fluorescence micrographs of HeLa cells transfected separately with pGFP-GST-RRRKR (panel a), pGFP-GST-TTF2 (1–80) (panel b), pGFP-GST-TTF2 (1–110) (panel c), pGFP-GST-TTF2 (80–151) (panel d), and pGFP-GST-TTF2 (110–151) (panel e).

proteins (Fig. 3A), excluding that the diffused fluorescence might derived by protein degradation. Plausibly, the cytoplasmic distribution of the fluorescent signal of GFP-TTF2 (1–50) could be the result of passive diffusion through the NPC, in addition to a less efficient nuclear retention due to the absence of the DNA-binding domain. The involvement of DNA recognition motifs in nuclear accumulation has been demonstrated for two thyroid transcription factors: in TTF1 a nuclear targeting sequence is located in the DNA recognition helix of the homeodomain and acts in concert with a basic NLS identified at the N-terminal end of the homeodomain. [31]; in the Pax 8 factor, the paired box that interacts with the DNA contributes to nuclear localization in addition to a second region located in the residual homeodomain [32].

In order to investigate the contribution to nuclear accumulation or retention of residues inside of the TTF2 FHD, we analyzed the cellular distribution of different fusion proteins expressing TTF2 fragments 1–151 mutated by deletions in the DNA-binding domain (Fig. 4, panel A). All constructs were analyzed by Western blotting using an anti-GFP antibody to ascertain the appropriate full length (Fig. 3B).

A strictly nuclear localization was observed when deletions at positions 64–65 (ΔMA), 75–76 (ΔTL), 95–96 (ΔKK), 116–117 (ΔRE) or 134–135 (ΔED) were introduced separately in the GFP-fusion proteins (Fig. 4B, b, d, e, g, h), whereas an additional cytoplasmic distribution of the fluorescent signal appeared by deletion of residues 72–73 (ΔRR), between helix 1 and 2 of the FHD and, less evident, by deletion of residues 102–103 (ΔRH), in helix 3 (Fig. 4B, c, f). The cytoplasmic distribution of the fluorescence signal expressed by ΔRR and ΔRH constructs might be due to a less efficient nuclear retention of the deleted fusion proteins, which are proximal to the passive limit of diffusion, or to a presence of a more complex NLS sequence at the N-terminal of FHD. In order to address this issue, we have made a construct that contains the TTF2 RRRKR motif cloned into a pGFP-GST vector, expressing a fusion protein larger than the passive diffusion limit through the NPC. As shown in Fig. 5 (panel a), the fusion protein was distributed exclusively into the nucleus, but compared to the nuclear localization of the GFP-TTF2 (1–376), the punctuated subnuclear accumulation disappeared. We further analyzed the expression of fusion proteins containing parts of the FHD. Expression in transfected HeLa cells of residues 1 to 80, 1 to 110, 80 to151 or 110 to 151, fused separately to GFP-GST, showed an exclusive nuclear distribution (Fig. 5, b, c, d, e). The punctuated pattern of nuclear accumulation disappeared in the distribution of fusion proteins deleted at the N-terminal of the FHD up to residue 110. These results demonstrated that the basic stretch RRRKR is a bona fide NLS, and that residues at the N-terminal FHD may contribute to nuclear retention.

The functional contribution to nuclear localization given by the N- and C-terminal boundary of the FHD has been demonstrated for FOXA2 and FOXF2 [23,24]. In FOXC1 and FOXO4, only the basic region conserved at the C-terminal end of the FHD has been demonstrated to represents a bona fide NLS [25,26]. The presence of nuclear determinants close to, or inside to, the DNA binding are often demonstrate in DNA and RNA binding proteins [33,34]. The presence of identical NLSs near or within the domain involved in DNA or RNA binding is less frequent. TTF2 is the only human FOX protein that contains identical basic stretches at the N- and C-terminal ends.

3.2. TTF2 binds importin α

Proteins carrying "classical" NLSs bind generally to the cytoplasmic receptors belonging to the importin α family. There are, however, many exceptions to this type of import, derived from different types of cytoplasm receptors that can recognize the NLS [10], or the NLS that bind directly to importin β, independently of importin α [35], or proteins that can translocate through the NPCs in the absence of any

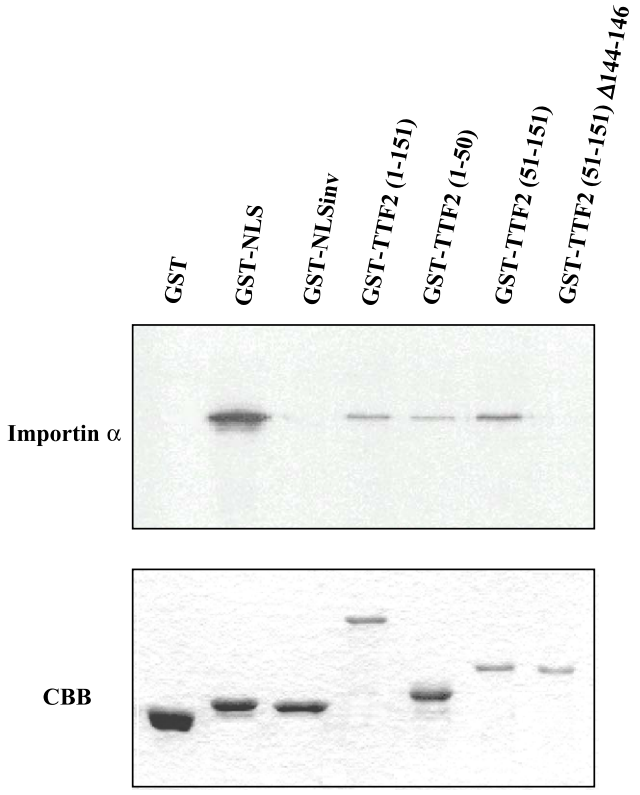


Fig. 6. Binding of importin α to TTF2. GST-TTF2 (1–151), GST-TTF2 (1–50), GST-TTF2 (51–151), GST-TTF2 (51–151) 144–146, GST-NLS or GST-NLSinv, fusion proteins immobilized on glutathione sepharose 4B, were incubated with in vitro translated importin α labeled with [^{35}S]-methionine. The proteins were separated by SDS-10% PAGE gel, and bound importin α was analyzed by fluorography (upper panel). An amount representing 1/20 of the beads incubated with GST fusion proteins, extensively washed, was resolved by SDS-PAGE and stained with Coomassie brilliant blue (lower panel, CBB staining).

soluble factors [36,37]. We tested the ability of a human importin α to bind the fragment of TTF2 containing both NLSs, or a single one, compared to the fragment mutated by deletions in both NLSs. We performed binding assays using in vitro translated importin α and recombinant GST-TTF2 fusion proteins expressed in bacterial cells. We used GST fusion proteins with SV40 Tag NLS (GST-NLS) or an inverse version of Tag NLS (GST-NLSinv) as positive and negative controls, respectively, of importin α binding. As shown in Fig. 6, importin α binds to GST-TTF2 (1–151), GST-TTF2 (1–50) and GST-TTF2 (51–151). Importin α binding was not detected using the GST-TTF2 (51–151) Δ 144–146 peptide that contains the FHD deleted of basic residues at either end. These results suggest that TTF2 nuclear import occurs via an importin α-dependent pathway and that at least one of the two NLSs is required for cytosolic receptor recognition.

In conclusion, we have identified the structural determinants for TTF2 nuclear localization. We have shown that TTF2 protein contains two identical bona fide NLSs at both ends of the DNA-binding domain that are required for importin α receptor binding. In addition, we have shown that residues inside the FHD may contribute to full accumulation of the TTF2 protein into the nucleus.

Acknowledgements

We thank Michael F. Rexach, Karsten Weis and Anna-Maria Frischauf for their generous gift of GST-NLS and GST-NLSinv constructs, pRSET-hSRP α plasmid and FKHL15 cDNA, respectively. This work was supported by grants from MIUR.

References

[1] K.S. Gajiwala, S.K. Burley, Winged helix proteins, Curr. Opin. Struck. Biol. 10 (2000) 110–116.
[2] E. Kaufmann, W. Knochel, Five years on the wings of fork head, Mech. Dev. 57 (1996) 3–20.
[3] K.H. Kaestner, W. Knochel, D.E. Martinez, Unified nomenclature for the winged helix/forkhead transcription factors, Genes Dev. 14 (2000) 142–146.
[4] E. Lai, K.L. Clark, S.K. Burley , J.E. Darnell Jr., Hepatocyte nuclear factor 3/fork head or "winged helix" proteins: a family of transcription factors of diverse biologic function, Proc. Natl. Acad. Sci. U. S. A. 90 1993, pp. 10421–10423.
[5] L. Crisponi, M. Deiana, A. Loi, F. Chiappe, M. Uda, P. Amaldi, L. Bisceglia, L. Zelante, R. Nagaraja, S. Porcu, M.S. Ristaldi, R. Marzella, M. Rocchi, M. Nicolino, A. Lienhardt-Roussie, A. Nivelon, A. Verloes, D. Schlessinger, P. Gasparini, D. Bonneau, A. Cao, G. Pilia, The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome, Nat. Genet. 27 (2001) 159–166.
[6] A.J. Mears, T. Jordan, F. Mirzayans, S. Dubois, T. Kume, M. Parlee, R. Ritch, B. Koop, W.L. Kuo, C. Collins, J. Marshall, D.B. Gould, W. Pearce, P. Carlsson, S. Enerback, J. Morissette, S. Bhattacharya, B. Hogan, V. Raymond, M.A. Walter, Mutations of the forkhead/winged-helix gene, FKHL7, in patients with Axenfeld–Rieger anomaly, Am. J. Hum. Genet. 63 (1998) 1316–1328.
[7] C.S. Lai, S.E. Fisher, J.A. Hurst, F. Vargha-Khadem, A.P. Monaco, A forkhead-domain gene is mutated in a severe speech and language disorder, Nature 413 (2001) 519–523.
[8] E.V. Semina, I. Brownell, H.A. Mintz-Hittner, J.C. Murray, M. Jamrich, Mutations in the human forkhead transcription factor FOXE3 associated with anterior segment ocular dysgenesis and cataracts, Hum. Mol. Genet. 10 (2001) 231–236.
[9] R.J. Clifton-Bligh, J.M. Wentworth, P. Heinz, M.S. Crisp, R. John, J.H. Lazarus, M. Ludgate, V.K. Chatterjee, Mutation of the gene encoding human TTF-2 associated with thyroid agenesis, cleft palate and choanal atresia, Nat. Genet. 19 (1998) 399–401.
[10] E.D. Schwoebel, M.S. Moore, The control of gene expression by regulated nuclear transport, Essays Biochem. 36 (2000) 105–113.
[11] I.W. Mattaj, L. Englmeier, Nucleocytoplasmic transport: the soluble phase, Annu. Rev. Biochem. 67 (1998) 265–306.
[12] D. Görlich, U. Kutay, Transport between the cell nucleus and the cytoplasm, Annu. Rev. Cell Dev. Biol. 15 (1999) 607–660.
[13] E.A. Nigg, Nucleocytoplasmic transport: signals, mechanisms and regulation, Nature 386 (1997) 779–786.
[14] D. Christophe, C. Christophe-Hobertus, B. Pichon, Nuclear targeting of proteins: how many different signals? Cell Signal. 12 (2000) 337–341.
[15] D.A. Jans, C.Y. Xiao, M.H. Lam, Nuclear targeting signal recognition: a key control point in nuclear transport? BioEssays 22 (2000) 532–544.
[16] M.G. Prieve, K.L. Guttridge, J. Munguia, M.L. Waterman, Differential importin-α recognition and nuclear transport by nuclear localization signals within the high-mobility-group DNA binding domains of lymphoid enhancer factor 1 and T-cell factor 1, Mol. Cell. Biol. 18 (1998) 4819–4832.
[17] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 1989.
[18] B.P. Chadwick, F. Obermayr, A. Frischauf, FKHL15, a new human member of the forkhead gene family located on chromosome 9q22, Genomics 41 (1997) 390–396.
[19] M. Rexach, G. Blobel, Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins, Cell 83 (1995) 683–692.
[20] K. Weis, I.W. Mattaj, A.I. Lamond, Identification of hSRP1 alpha as a functional receptor for nuclear localization sequences, Science 268 (1995) 1049–1053.
[21] M.G. Romanelli, C. Morandi, Importin α binds to an unusual bipartite nuclear localization signal in the heterogeneous ribonucleoprotein type I, Eur. J. Biochem. 269 (2002) 2727–2734.
[22] P.E. Macchia, M. Mattei, P. Lapi, G. Fenzi, R. Di Lauro, Cloning, chromosomal localization and identification of polymorphisms in the human thyroid transcription factor 2 gene (TTF2), Biochimie 81 (1999) 433–440.
[23] J.D. Licht, M.J. Grossel, J. Figge, U.M. Hansen, Drosophila Krüppel protein is a transcriptional repressor, Nature 346 (1990) 76–79.
[24] S. Saha, J.M. Brickman, N. Lehming, M. Ptashane, New eukariotic transcriptional repressor, Nature 363 (1993) 648–652.
[25] W. Hanna-Rose, U. Hansen, Active repression mechanism of eukaryotic transcription repressors, Trends Genet. 12 (1996) 229–234.
[26] J. Sutton, R. Costa, M. Klug, L. Field, D. Xu, D.A. Largaespada, C.F. Fletcher, N.A. Jekins, N.G. Copeland, M. Klemsz, R. Hromas, Genesis, a winger helix transcriptional repressor with expression restricted to embryonic stem cells, J. Biol. Chem. 271 (1996) 23126–23133.
[27] X. Qian, R.H. Costa, Analysis of hepatocyte nuclear factor-3β protein domains required for transcriptional activation and nuclear targeting, Nucleic Acids Res. 23 (1995) 1184–1191.
[28] M. Hellqvist, M. Mahlappu, A. Blixt, S. Enerback, P. Carlsson, The human forkhead protein FREAC-2 contains two functionally redun-

dant activation domains and interacts with TBP and TFIIB, J. Biol. Chem. 273 (1998) 23335–23343.
[29] F.B. Berry, R.A. Saleem, M.A. Walter, FOXC1 transcriptional regulation is mediated by N- and C-terminal activation domains and contains a phosphorylated transcriptional inhibitory domain, J. Biol. Chem. 277 (2002) 10292–10297.
[30] A.M. Brownawell, G.J. Kops, I.G. Macara, B.M. Burgering, Inhibition of nuclear import by protein kinase B (Akt) regulates the subcellular distribution and activity of the forkhead transcription factor AFX, Mol. Cell. Biol. 21 (2001) 3534–3546.
[31] C. Christophe-Hobertus, V. Duquesne, B. Pichon, P.P. Roger, D. Christophe, Critical residues of the homeodomain involved in contacting DNA bases also specify the nuclear accumulation of thyroid transcription factor-1, Eur. J. Biochem. 265 (1999) 491–497.
[32] A. Poleev, O. Okladnova, A.M. Musti, S. Schneider, B. Royer-Pokora, D. Plachov, Determination of functional domains of the human transcription factor PAX8 responsible for its nuclear localization and transactivating potential, Eur. J. Biochem. 247 (1997) 860–869.
[33] E.C. LaCasse, Y.A. Lefebvre, Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins, Nucleic Acids Res. 23 (1995) 1647–1656.
[34] M. Cokol, R. Nair, B. Rost, Finding nuclear localization signals, EMBO Rep. 1 (2000) 411–4151.
[35] S. Jakel, D. Gorlich, Importin β, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells, EMBO J. 17 (1998) 4491–4502.
[36] F. Yokoya, N. Imamoto, T. Tachibana, Y. Yoneda, β-Catenin can be transported into the nucleus in a Ran-unassisted manner, Mol. Biol. Cell 10 (1999) 1119–1131.
[37] F. Fagotto, U. Gluck, B.M. Gumbiner, Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin, Curr. Biol. 8 (1998) 181–190.